Recommendations from the Nevada Accountability Advisory Committee for the Revised School Performance Framework

A Report Submitted to the Nevada Department of Education by Elena Diaz-Bilello and Chris Domaleski

National Center for the Improvement of Educational Assessment, Inc.

Center for Assessment

Executive Summary

The National Center for the Improvement of Educational Assessment (Center for Assessment) worked with the Nevada Department of Education (NDE) to explore and document recommended improvements to the Nevada School Performance Framework (NSPF). A central part of the process involved convening a broad based committee of state education and community leaders – the Accountability Advisory Committee (AAC). This process culminated in this report which documents the process and set of recommendations made by the AAC and illustrate strategies for NDE's consideration when implementing AAC design recommendations.

The Center for Assessment, NDE, and AAC met in person monthly starting in January, 2015 through May, 2015. The objectives for the meetings were to:

1. Define accountability goals aligned with CCR expectations.
2. Define design principles to help operationalize the selection and design of indicators to meet accountability goals.
3. Ensure alignment of or redefine current metrics and indicators to support CCR expectations for all students.
4. Ensure that performance standards set to designate ratings can be connected to CCR goals for all students.

The AAC addressed the first three purposes noted above. However, the fourth area concerning the setting of performance standards on overall ratings was necessarily deferred to the summer of 2016 due to an unforeseen lack of state assessment data in 2015. The outcomes for each of these areas are summarized below.

Goals

Clear goal statements serve to direct and help prioritize decisions about the framework. For this reason, the AAC devoted considerable time to develop and refine the goals. The following statements represent the prominent consensus decision of the AAC with respect to system goals.

- The system must provide clear, actionable information to help districts and schools evaluate and improve the effectiveness of their programs
- The system must accurately classify districts and schools with respect to performance standards in order to inform the public and policy makers
- Ratings should reflect and help promote
    - Improved academic achievement, especially with respect to growth and equity of outcomes
    - Progress toward and attainment of post-secondary readiness
    - Positive school climate

Design Principles

Another set of important guiding decisions to inform the development of the framework is the design principles. If the goals represent the intended destination on a roadmap, the design principles serve to establish the nature and manner of the route. The principles listed below were established by the AAC and are intended to make clear the features and conditions that will

characterize a framework that successfully supports the goals discussed in the previous section. The design principles also serve as a basis to evaluate that the system is working as intended.

- Growth should be weighted more heavily than status performance
- The full range of accountability performance should be accessible to schools of all types, including those that serve at risk students
- Outcomes should be consistent and comparable from school to school
- The outcomes should reflect and reward reduction in performance gaps
- Where reasonable and without sacrificing technically defensibility, the model should be as simple as possible to promote understanding
- Overall, indicators should be compensatory but incorporate established thresholds and weighting decisions

Framework Components

After establishing consensus on the goals and design principles for the system, the next set of considerations addressed by the AAC focused on the NSPF system components and indicators. These are briefly summarized below.

Status and Growth Measures of Achievement

The AAC did not make any suggestions to re-consider the current approaches used to evaluate the metrics used for status and growth in the NSPF. Although the approaches to evaluate academic status and growth on the summative state assessments will remain the same, the AAC unanimously supported a recommendation to include performance achieved on the science tests administered state-wide as an integral part of the status component.

Equity

The AAC affirmed that equity should be reflected prominently in the accountability framework. Broadly, equity refers to indicators that reflect the extent to which a school is improving academic outcomes for the most at risk students.

Overall, the AAC endorsed the current calculation methodology, which involves evaluating group MGP with respect to AGP. However, the discussion focused heavily on how best to define the equity group(s). A majority view emerged in support of evaluating gaps relative to membership in academic performance groups. All AAC members supported an approach informed by additional research that considered a solution with respect to the following principles.

- Ensure that traditionally at-risk groups are represented in the equity groups, particularly racial and ethnic minorities.
- Implement measures to ensure that schools may not receive favorable outcomes if the performance of any one racial/ ethnic subgroup falls below an established threshold
- Continue to prominently report the performance of ALL subgroups as is the current practice

College and Career Readiness

The AAC reviewed each of the four CCR indicators currently in the framework and provided several recommendations and considerations for re-evaluating this component.

Recommendations for modifying current indicators were:

- NDE should consider revising the percentage of advanced diploma earners indicator to include all high school diplomas earned by students.
- Substitute the percentage of students participating on ACT or SAT tests with the percentage of students meeting college-ready benchmarks on either test

The AAC noted that all of the current indicators only focus on college readiness and that a career readiness focus is missing in this component. Some suggestions provided by the AAC members included:

- Explore the feasibility of using some of the non-academic skills outlined in David Conley's framework.
- Offer other CCR indicator options to account for inequities in resources across schools

School Climate

Measures of school climate represent an important recommended addition to the NSPF. The AAC urged that school climate should not be punitive, exerting a moderate to minimal impact on a school's overall rating. Three categories for identifying school climate indicators emerged.

- The first category would consist of indicators common to all school such as discipline data, and attendance data.
- The second category would consist of indicators where schools have flexibility in selection. This provides an opportunity for schools to select indicators that would showcase the good work they do to improve school quality for teachers, students, parents, and the larger community.
- The third area serves as an optional "test-site" where schools can choose to collect and report results from these indicators but not use the information for accountability.

Conclusions and Next Steps

Moving forward, the following two phase process is recommended to produce the revised operational system. This process takes into account the challenge presented by the unexpected loss of summative assessment data in 2015.

- Phase I: Analyze Data and Develop Model Revisions (Summer 2015 – Fall 2015)
- Phase II: Standard Setting (Summer 2016)

The first set of complete statewide data will not be available until following the 2016 administration. For this reason, we recommend delaying standard setting to the summer of 2016 when the impact of academic achievement measures can be evaluated and incorporated in the standard setting process. This will enable the process to utilize school profiles that accurately portray performance. Moreover, it will be prudent to conduct an additional review in the summer of 2017 before confirming the standards.

Introduction

This report provided by the National Center for the Improvement of Educational Assessment (Center for Assessment) to the Nevada Department of Education (NDE) serves two purposes:

- To document the process and set of recommendations made by the Accountability Advisory Committee (AAC) for re-designing the Nevada School Performance Frameworks (NSPF).

- To illustrate different strategies for NDE's consideration when exploring and implementing AAC design recommendations.

The NSPF re-design work was initiated based on recognition from the NDE leadership that although the current indicators in the NSPF provide a useful basis for evaluating the quality of academic performance across schools, they questioned the extent to which the current rating system provided adequate signals to schools and the broader public that students were meeting college and career readiness (CCR) expectations. To ensure that the performance standards set in the framework provided credible information about CCR, the NDE launched a process to re-evaluate the current standards set for rating all schools. Part of this process included the formation of the AAC comprised of community-based, higher education and district leaders to provide input on the standard setting work.

Although the impetus for the re-design work came largely out of concerns to ensure that the overall ratings on the NSPF can be connected to larger CCR goals, NDE recognized that this work would require moving beyond evaluating the standards set and would also require re-examining the set of underlying indicators that provide the foundation for the overall ratings. To meet this re-design objective, the NDE contracted with the Center for Assessment to design a process in partnership with NDE staff that would guide AAC members to develop recommendations in the following four areas:

1. Defining accountability goals aligned with CCR expectations.

2. Defining design principles to help operationalize the selection and design of indicators to meet accountability goals.

3. Ensure alignment of or redefine current metrics and indicators to support CCR expectations for all students.

4. Ensure that performance standards set to designate ratings can be connected to CCR goals for all students.

The AAC work to date has largely addressed and culminated with a set of recommendations for the first three areas noted above. The fourth area concerning the setting of performance standards on overall ratings will be addressed by the AAC in the summer of 2016.

In addition to guiding the AAC's work with NDE staff, the Center for Assessment agreed to develop this report to document the process used to gather input from AAC members and the resulting recommendations. This report begins with an overview of the AAC's role and provides a summary of the set of topics and different approaches used to gather input and recommendations from all members. Following the summary of the process used to guide AAC recommendations, we document the goals and design principles defined by consensus from all AAC members. After documenting the set of goals and design principles recommended for the next version of the NSPF, we then capture the set of input and recommendations made by AAC members to better align and/or redefine current metrics and indicators to support CCR expectations for all students. To illustrate possible approaches for implementing the AAC's recommendation and input, we outline several approaches for NDE's consideration. This report concludes with a brief overview of next steps to operationalize the model.

The Role of the AAC

Forming the Accountability Advisory Committee

To move forward with the process of re-designing the NSPF, the Nevada Department of Education (NDE) convened the AAC comprised of leaders from key stakeholder groups. In consultation with the Superintendent of Public Instruction, Dale A.R. Erquiaga, Deputy Superintendent for Student Achievement, Steve Canavero, contacted leaders of important Nevada constituencies and invited them to participate in providing recommendations for improving the current NSPF. The stakeholder groups invited included leaders of school district and the State Public Charter School Authority, various student populations, parents, educator organizations (i.e., unions), the Nevada System of Higher Education, the business community, non-profit organizations, and on-site educators. The complete list of AAC members is located in Attachment A.

At the start of each meeting, AAC members were reminded of their charge to:

- Determine design priorities for Nevada's school accountability system.
- Establish performance expectations for the accountability system.

Since the AAC serves as an advisory body, the recommendations provided and information captured in this document are not intended to reflect policy for designing the frameworks. Rather, these recommendations will be carefully considered by other stakeholder groups including the NDE leadership and may be accepted to inform policy.

Beginning in January through May, the AAC convened once a month for a total of 5 meetings. Table 1 presents an outline of topics addressed at each meeting.

Table 1. Topics addressed in AAC Meetings

Meeting Date	Topics Addressed
January 20, 2015	• Setting the ground work for designing an accountability system
February 27, 2015	• Identifying accountability goals and design principles for the NSPF • Coming to consensus on priority goals and design principles
March 18, 2015	• Considering design illustrations for each indicator category in the NSPF: performance gaps
April 22, 2015	• Considering design illustrations for each indicator category in the NSPF: college and career readiness (CCR). • Closure on performance gaps
May 22, 2015	• Considering design illustrations for each indicator category in the NSPF: school climate. • Re-evaluating the weights assigned to each indicatory category in the NSPF

In January and February, the AAC members worked on achieving consensus on key design priorities and goals to help determine which components and indicators in the NSPF needed to be modified or re-evaluated. Based on consensus achieved on the key design priorities and goals, the next set of topics from February through May focused on the key components and indicators in the NSPF that required a re-evaluation to ensure that these areas met design priorities and CCR expectations. As indicated by Table 2, these topics centered on re-evaluating: subgroups used to monitor the academic progress made by all students, the current set of CCR indicators, and the current weighting scheme in the framework. The AAC also considered the inclusion of a new school climate component for the NSPF. There were two components where AAC members did not recommend new metrics to use:

- Ensuring that academic achievement continues to report the percent proficient by subject area on the NSPF.

- Ensuring that academic growth continues to report the median growth percentiles (MGPs) on the NSPF.

Although the metrics used to report academic achievement and academic growth will remain the same, the AAC recommended that the percent proficient and above for science be included under academic achievement. Additionally, the AAC voiced concern about the assessment transition's impact on reporting academic growth for high schools. These two issues will be discussed in more detail in a later section of this report.

The AAC will meet again to review the model and to provide input on setting performance standards for schools. These standards will be used to determine overall school ratings aligned to CCR expectations. During that meeting, AAC members will be asked to provide input to define expected performance profiles of schools falling under each rating

category. These profiles will then be used by NDE to help inform the setting of new standards for assigning final ratings to all schools.

Process Used to Solicit Input and Recommendations

To welcome and encourage a diversity of opinions, a consensus-based approach to developing recommendations was achieved, with dissenting views noted for the record. For each AAC meeting, the Center for Assessment in partnership with NDE presented different design options and considerations in reference to the set of topics noted in Table 1. AAC members were then asked to provide their recommendations by submitting input through facilitated activities and different feedback forms. Table 2 presents an inventory of methods and the process used to gather input from AAC members at each meeting.

Table 2. Inventory of methods and process used to collect input from AAC members

Meeting Date	Methods
Pre-meeting	• The Center for Assessment administered a pre-meeting survey to all participants to gather initial data about stakeholder understanding and perceptions about NSPF. • Data from the survey was used to determine the extent to which stakeholders would like to see changes made to each component of the framework as well as the overall rating system established.
January 20, 2015	• Small group activity organized to identify common themes regarding perspectives about school accountability from the pre-meeting survey. • Small group activity organized to evaluate range of preferences on school quality characteristics to shape goals and design principles for the NSPF.
February 27, 2015	• Based on input received from small group activities from January, initial set of accountability goals and design principles for the NSPF provided to AAC members on chart paper. Small group activity organized to gather input from groups to prioritize goals and selected design principles. • Common set of goals and design principles identified across small groups to form the basis for establishing the priority goals and design principles for the NSPF.
March 18, 2015	• Shared approaches for how and which subgroups are evaluated in other states to provide context for possible design choices. • Presented analyses on considering different subgroup configurations: evaluating performance of lowest achieving group, all federally defined groups, and current groups used in the NSPF. • Approaches and analyses used to solicit preliminary input on subgroup configurations.
April 22, 2015	• Small activity organized to capture small group input on

	feedback forms to define subgroups and re-consider the set of CCR indicators in the NSPF. • Small group feedback shared with whole group to determine common areas to define recommendations and identify differences.
May 22, 2015	• Small activity organized to capture small group input on feedback forms to consider the use of a School Climate indicator and to reconsider the current weighting of indicators in the NSPF. • Small group feedback shared with whole group to determine common areas to define recommendations and identify differences.

For each of the major areas requiring careful consideration, we structured and facilitated different activities with AAC members to ensure that consensus could be established through a three-step process:

1. Capturing individual input.
2. Ensuring that individual input was used to inform small group contributions and consensus-making at the small-group level.
3. Using small-group consensus to identify common whole-group recommendations and to note any key differences.

During each of the AAC meeting, the Center for Assessment captured meeting notes to document recommendations reflecting the majority of perspectives shared across whole group members, as well as to highlight areas where one or two members dissented with the larger group. The meeting notes served as a running record throughout the January to May period to ensure that all members agreed with the set of key recommendations reached and to ensure that all critical decision points and issues had been adequately framed. In this report, all of the recommendations and issues reflected include all updates and edits provided by AAC members. In the next section, we address the goals and design principles selected by the AAC members to help guide discussions and recommendations for the NSPF.

Establishing Goals and Design Principles

Goals

The design of any accountability system must be guided by explicit goals and intended outcomes. Clear goal statements serve to direct and help prioritize decisions about the framework. For this reason, the AAC devoted considerable time to develop and refine the goals early in the process. The following statements represent the prominent consensus decision of the AAC with respect to system goals.

- **The system must provide clear, actionable information to help districts and schools evaluate and improve the effectiveness of their programs.**

This statement reflects a shared recognition that the primary purpose of the system is to provide useful feedback to educators to gauge the success of the school's efforts to improve student achievement. Moreover, such information guides decisions about allocation of resources, such as where to focus professional development or which programs should be expanded or reduced. In order to realize this goal, ACC stressed that accountability information must be clear, specific, and useful to education leaders for evaluation, planning, and decision-making.

- **Accurately classify districts and schools with respect to performance standards in order to inform the public and policy makers.**

  The AAC also recognized that a central goal of the state accountability system is to provide information to a variety of stakeholders to better understand the performance of Nevada's schools. Policy makers represent a key stakeholder group that relies on information from the accountability system to evaluate the efficacy of current initiatives and potentially to inform new policy. Additionally, members of the community deserve current and accurate information to inform educational decisions. More broadly, as a matter of principle, the accountability system represents an important part of the understood responsibility that is concomitant with the public investment and trust placed in schools.

- **Ratings should reflect and help promote:**
  - **Improved academic achievement, especially with respect to growth and equity of outcomes.**
  - **Progress toward and attainment of post-secondary readiness**
  - **Positive school climate**

  The final, multifaceted goal statement reflects both the dimensions of quality on which school ratings should be based and the aspiration that inclusion of these elements in the system will help incentivize progress. With respect to the latter claim, there was a sense shared by many AAC members that 'what is measured is what matters.' While other AAC members were less confident that inclusion of selected elements (alone) will appreciably motivate performance, there was, notwithstanding, unanimous agreement to prioritize the components listed in this goal statement as important markers of school quality. More details about these components will be provided in subsequent sections of this report.

Design Principles

Another set of important guiding decisions to inform the development of the framework is the design principles. If the goals represent the intended destination on a roadmap, the design principles serve to establish the nature and manner of the route. The principles listed below were established by the AAC and are intended to make clear the features and conditions that will characterize a framework that successfully supports the goals discussed in the previous section. The design principles also serve as a basis to evaluate that the system is working as intende

- **Growth should be weighted more heavily than status performance**

  The AAC affirmed that academic growth (i.e., student growth percentile) should have more influence on the final rating than status (i.e., percent proficient). This principle emerged in recognition of the fact that status is often strongly connected to the population of students that a school serves, whereas growth better reflects the contributions that teachers and leaders make to student learning over the course of the academic year. To be clear, status is important as the attainment of established performance standards remains a fixed goal. However, the preferred measure of academic performance for school accountability includes both status and growth to a meaningful standard, where the latter more heavily influences the final outcome. For this reason, several AAC members expressed concerns about evaluating student outcomes for high schools where growth will not be computed in the foreseeable future.

- **The full range of accountability performance should be accessible to schools of all types, including those that serve at risk students.**

  As an extension of the previous principle, the AAC noted that the accountability system must be designed such that schools serving at risk students are not systematically disadvantaged. For example, if the resulting model shows that schools in the most economically disadvantaged communities all receive the lowest ratings, this may be more likely an artifact of a poorly designed model than a true indication of school quality[1]. For this reason, the resulting model must be evaluated to determine if the full range of outcomes are available to schools of all types.

  However, the AAC indicated that schools serving a high population of at-risk students, such as alternative education schools, should be evaluated using different indicators or a different framework altogether.[2] For these schools, stringent criteria should be established to ensure that a school meets specific criteria (e.g., 95% of students are under-credit and over-age students).

  In addition to recognizing that alternative schools may need to be evaluated using a separate accountability framework, the AAC also noted that contextual factors that could impact the ratings of a school should be considered but not "baked" into the model. The implication of this recommendation is that direct adjustments or covariates will not be used to evaluate "similar" schools on indicators such as growth and achievement, but rather these factors can be considered in the following ways:

  - Ensuring that an optimal n size is set for each indicator.

[1] In other words, this finding would raise a flag and should instigate deeper investigate work that may include on-site reviews to determine whether there are other factors that may be contributing to systematic low performance or whether this rating is potentially warranted.

[2] Per the NDE, the development of an alternative education framework is being considered for alternative education schools.

- Using "triggers" or business rules to ensure a school with large performance gaps do not qualify for the highest rating.
- Consider the possibility of using a review process that would draw upon qualitative approaches to better evaluate the impact of these contextual factors on performance outcomes for individual cases. Findings from on-site reviews or interviews with school personnel may assist in unique cases where alternative indicators outside of the framework may need to be consulted to clarify the performance picture of a given school.

- **Outcomes should be consistent and comparable from school to school**

  The AAC discussed at length, the tradeoffs between standardization and flexibility as a principle of accountability design. Ultimately, the AAC privileged an approach in which measures are reasonably uniform to produce results that can compare the performance of one school in contrast to another. Ensuring comparability of scores was deemed by the AAC as critical to achieve the stated goals of evaluating effectiveness and providing meaningful information to stakeholders.

- **The outcomes should reflect and reward reduction in performance gaps**

  The AAC agreed that promoting improved academic achievement for low-performing students is an important focal area for Nevada schools. For this reason, the system must meaningfully distinguish between schools that are more or less successful in reducing achievement gaps. Schools that successfully reduce gaps should be recognized and rewarded to both incentivize other schools to reduce gaps and to shed light on promising practices and strategies at those schools.

- **Where reasonable and without sacrificing technically defensibility, the model should be as simple as possible to promote understanding**

  In order to ensure the model achieves the stated goal of providing 'actionable' information, the system should be clear, transparent, and avoid unnecessary complication. Stated another way, if there are multiple design approaches under consideration, each of relatively equal merit, the more straightforward approach should be selected.

- **Overall, indicators should be compensatory but incorporate established thresholds and weighting decisions**

  The AAC recognized that a quality school may be high performing in many but not all areas. Given that the accountability system includes multiple indicators, it is reasonable to design the system such that higher performance in one area may offset lower performance in another area. In some cases, as noted in the first design principle (i.e., growth should be weighted more heavily than status performance**)**
  , the compensatory nature of the principle may favor one component over another. In other cases, there may be limits established on the degree of offset permissible to achieve a certain outcome. This use of thresholds or 'triggers' will be discussed in more detail in a subsequent section of this document.

System Components

Having established consensus on the goals and design principles for the system, the next set of considerations addressed by the AAC focused on the NSPF system components. All elementary and middle schools in the NSPF have the following system components present in the framework:

1. Growth Measure of Achievement – **40 points**
2. Status Measure of Achievement – **30 points**
3. Reductions in Achievement Gaps – **20 points**
4. Other Indicators as defined by average daily attendance – **10 points**

For high schools, the system components with total points assigned are as follows:

1. Status/Growth Measure of Achievement – **30 points**
2. Reductions in Achievement Gaps – **10 points**
3. Graduation Measures – **30 points**
4. College and Career Readiness – **16 points**
5. Other Indicators as defined by average daily attendance and the percent of 9th grade students who have earned at least five credits by the end of 9th grade. – **14 points**

For all schools, the points earned are summed across all components to derive an index or composite score. This index score is then used to determine the overall "star" performance plan assignment or rating.

This section, beginning with the updates made by the AAC to the Status and Growth Measures of Achievement component, documents the revisions made to two other components (CCR and performance gaps) and the overall weighting of points assigned to each component. We also outline the recommendation for adding a new School Climate component to the framework which would replace the "Other Indicators" component in the current framework. For each component where the AAC provided general parameters to modify the current design, we include possible design illustrations for NDE's consideration that would meet those parameters and subsequently also align with their supported goals for the accountability system and design principles for the NSPF.

Status and Growth Measures of Achievement

As indicated earlier in this report, the AAC did not make any suggestions[3] to re-consider the current approaches used to evaluate the metrics used for status and growth in the NSPF. Although the approaches to evaluate academic status and growth on the summative state assessments will remain the same, the AAC unanimously supported a recommendation to

[3] However, it is important to note here that these technical considerations were not part of the purview of the AAC. Although the AAC did not provide alternative suggestions to the current methodology used to evaluate academic status and achievement, other group such as the TAC may potentially do so.

include performance achieved on the science tests administered state-wide as an integral part of the Status component.

Recommendation for Including Science in the NSPF

During the May AAC meeting a State Board Education member, Mark Newburn, presented a case for including the science assessments in the framework which included:

- Not incorporating science in the frameworks has led schools to not place as much value, resources and attention to the sciences as ELA and Math.

- Decreased attention and focus on building appreciation for the sciences is highly problematic in that the education system is not producing enough students to fill the growing demand for individuals to fill Science, Technology, Engineering and Mathematics (STEM) and STEM-related occupations.

- STEM lacks diversity and would benefit greatly from encouraging more students of color as well as female students to enter the STEM pipeline. Additionally, although generating student interest in STEM fields should ideally begin in elementary schools; many elementary schools are not offering strong science programs.

AAC members agreed with the points made by the State Board member and also noted that there is an uneven focus and investment in sciences across schools. According to AAC members, including science in the frameworks would signal that providing increased access for all students to the sciences is valued by the state. To also demonstrate K-12 support for the governor's STEM initiative as well as to increase preparedness of students to move into these growing professions, the AAC recommended that science should be included as part of the status measures of achievement in the framework. AAC members noted that since fewer grades are currently tested in science relative to English language arts (ELA) and math, this content area should not be given the same weight as the ELA and math for the next version of the NSPF. The AAC recognized that growth cannot be computed for science at time but should additional science assessments be introduced in other grades, then a growth measure may be computed at a later time.

Considerations for Academic Growth

Concern was articulated by several AAC members about the lack of growth measures available for high schools beginning in the 2014-2015 school year. Although no specific recommendations were made about suitable substitutes for considering growth due to the absence of state tests for grades 9 and 10, at least 5 AAC members noted that this was a concern since the growth measures offer a different perspective of evaluating student learning relative to demonstrating proficiency on the academic achievement indicators. Should NDE decide to factor in growth measures for high school, this would require exploring additional measures administered state-wide outside of the current NSPF. Alternatively if districts have measures that could be used to evaluate student growth, these measures could potentially be explored as candidates. However, we note that this alternative pathway would conflict with the design principle to ensure that academic performance is evaluated using the same set of assessments.

Design Illustrations

Design illustrations are limited to just the inclusion of status since a discussion on alternative approaches to evaluating growth was not addressed as part of the AAC agenda[4]. However as indicated in the previous section, illustrations for growth are not presented in this section since analyses will need to be performed and input solicited from other technical groups to determine the suitability of using other assessments (e.g., ACT Aspire) for accountability.

Incorporating Science into Status

One illustration provided by one group of AAC members for incorporating science for status achievement in elementary and middle schools is provided in Table 3. Although the points awarded are evenly distributed for status, since there are no growth scores available for science, reading and math are still emphasized in the frameworks over the sciences. However, for high school, since growth will likely not be reported for some time at that level, a lower weight would need to be attributed to science for status achievement to keep consistent with the recommendation that science cannot carry the same weight as reading and math on the framework.

Table 3. Illustration of distributing points to science based on input from 1 group of AAC members

Elementary Status	Math	Reading	Science
Current NSPF status points	15	15	
Possible re-distribution of status points	10	10	10

Another possibility for NDE to consider is to re-distribute the total points assigned to status relative to the proportion of test takers in each content area state-wide. We draw on a simple example to illustrate how this can be carried out:

1. There are 1,000 students total (duplicated counts) of students taking either a math, reading, or science test.
2. Out of the 1,000 students, 400 students have math scores, 450 students have ELA scores and 150 students have science scores.
3. The proportion of total students with scores in each area is as follows: 40 percent for math, 45 percent for ELA, and 15 percent for science.

Table 4 presents the redistributed points for the framework relative to the proportion of test-takers with scores in each content area. The redistributed points are derived by multiplying the total points assigned to status by the percentage of test takers with scores in each content

[4] It is important to note here that AAC members were in agreement that although the growth measure is complex, they valued the information provided by the growth percentiles and did not want to see adjustments made with the percentiles.

area. For example, the six science points in Table 4 was calculated by taking the total points possible for status (30) and multiplying that by the proportion of test takers (15 percent or .15).

Table 4. Illustration of distributing points to science based on proportion of test-takers

Elementary Status	Math	Reading	Science
Current NSPF status points	15	15	
Possible re-distribution of status points	12	13 (rounded down)	5 (rounded up)

Equity

As noted previously in this document, the AAC affirmed that equity should be reflected prominently in the accountability framework. Broadly, equity refers to indicators that reflect the extent to which a school is improving academic outcomes for the most at risk students. Nevada's current school performance framework addresses equity by tracking the reduction in achievement gap for students with an Individualized Education Program (IEP), students who are English Language Learners (ELL) and students who qualify for a Free or Reduced Priced Lunch (FRL). In the elementary and middle schools, the reduction in achievement gap is measured by the percent of students meeting the Adequate Growth Percentile (AGP) target based on the median SGP. In high school, the reduction in achievement gap is measured by subtracting the performance of the percent proficient and above achieved by each of the focal groups (i.e., ELL, IEP, and FRL students) from the percent proficient and above achieved by all students.

Equity indicators were discussed more than any other topic over the course of the AAC meetings, which reflects the importance that members place on equity and the level of concern members exhibited to ensure the system 'got it right.' Overall, the AAC endorsed the current calculation methodology, which involves evaluating group MGP with respect to AGP. However, the discussion focused heavily on how best to define the equity group(s). Three general approaches emerged as candidates and the AAC carefully evaluated the pros and cons of each.

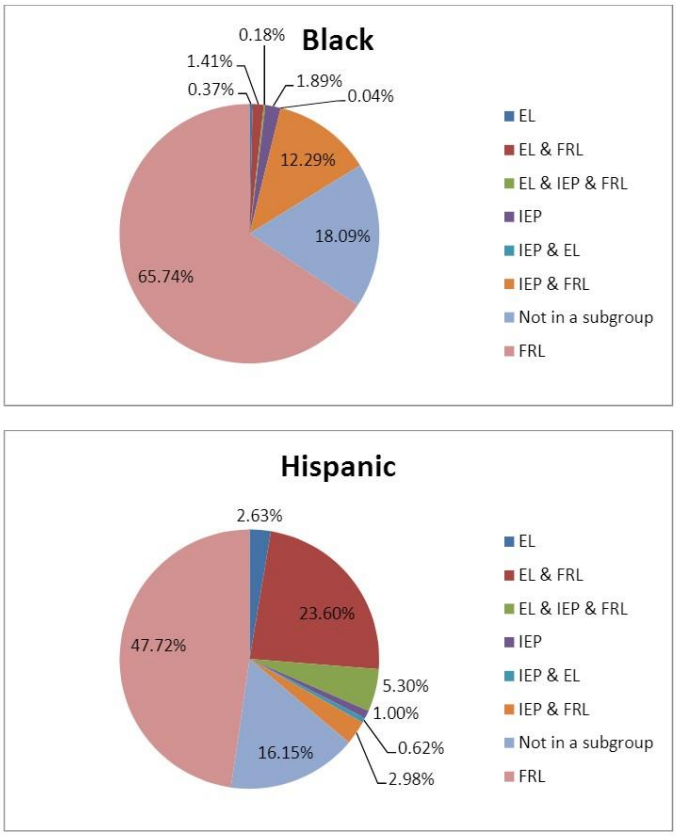
1. **Continue with the current NSPF approach.** As noted, the current approach establishes IEP, ELL, and FRL as the equity groups.
2. **Use all sub-groups.** Not unlike NCLB, this approach is based on evaluating performance of <u>all</u> subgroups <u>separately</u> to include groups defined by race/ ethnicity in addition to IEP, ELL, and FRL.
3. **Establish a performance based consolidated subgroup.** This approach involves creating a single subgroup based on a criterion or norm-referenced performance threshold.

In discussing the strengths and limitation of each approach, the AAC recognized that methods one and two offered the advantage of explicitly identifying specific, priority equity groups. In particular, the AAC advised that the performance of African American students and Hispanic students must be tracked since these groups continue to lag in performance relative to

all other subgroups. Moreover, this approach was seen as clear and straightforward to understand to some members, partly because it is familiar to many stakeholders. However, a disadvantage of options one and two is that students often appear in multiple groups (e.g. a student who is both FRL and ELL) which will likely lead to an undesirable outcome of systematically penalizing schools with more diverse populations (see Darling-Hammond, 2006; Novak & Fuller). As indicated earlier in the design principles section, this type of result would violate the second design principle. Additionally, for option two in particular, breaking the equity groups into many small groups will lead to many more schools not reporting equity performance data because of low n-sizes.

It is worth noting that the Nevada Department of Education presented data to the AAC revealing that African American and Hispanic students are prominently represented in the three groups (i.e. IEP, FRL, and ELL) already included in the frameworks. For example, Figure 1 below shows the percentage of African American or Black students included in an NSPF achievement gap subgroup and the percentage of Hispanic students assigned and not assigned to a subgroup currently. The results show that 82 percent of African American students and 84 percent of Hispanic students are current in an NSPF subgroup.

Figure 1. Percentage of Black students and Hispanic Students Assigned to an NSPF Achievement Gap Group.



In considering the third method, the creation of a performance based consolidated subgroup, the AAC recognized that this approach offered the advantage of focusing on all students in need of academic support regardless of the student's group membership. Moreover, many AAC members acknowledged that this mitigates the issue of school underrepresentation due to low n- size and is straightforward to understand and implement. However, two AAC members identified that a concern associated with this approach is the lack of explicitly identifying specific subgroups, particularly racial/ethnic groups, which is a central emphasis of equity. Most members did not share the same concern since they felt attention to specific subgroup performance could be addressed by building in "triggers" or business rules that prevent a school from either attaining a higher overall rating or higher points on this indicator if 1 or more subgroups exhibits low performance.

Recommendations

Although full agreement was not achieved on this issue, ultimately *the AAC noted a majority preference for evaluating gaps relative to membership in academic performance groups.* Regardless of position, all AAC members supported an approach informed by additional research that considered a solution with respect to the following principles.

- Ensure that traditionally at-risk groups are represented in the equity groups, particularly racial and ethnic minorities.
- Implement measures to ensure that schools may not receive favorable outcomes if the performance of any one racial/ ethnic subgroup falls below an established threshold (i.e. 'conjunctive trigger,' discussed later in this document).
- Continue to prominently report the performance of ALL subgroups as is the current practice.

Design Illustrations for Equity

In this section, we discuss two dimensions to implementing the equity recommendations that emerged from the AAC discussions: group definition and identifying the metric.

Equity Group Definition: Criterion or Norm Referenced

One approach is to establish an equity group based on a specific, static performance criterion on the state assessments. The most reasonable candidate for that criterion is the performance standard associated with College and Career Ready (CCR) or on-track to CCR, which is level three and higher on the Smarter-Balanced Assessment. Specifically, students who score below level 3 in the prior year are considered members of the consolidated equity group in the present year.

An advantage of this approach is that equity group membership is pegged to a meaningful threshold rooted in state policy. Using below proficient or not on track to achieve CCR as the equity focus is intuitively reasonable and provides a straightforward basis for schools to plan support initiatives. A potential drawback is that the subgroup size is likely to vary widely from

school to school, with some schools identifying many or most students in the equity group and other schools identifying relatively few.

Another approach is to define the equity group based on a performance threshold associated with a point in the distribution to yield a subgroup size that meets a specified target. For example, some states have used the first quartile – the point in the distribution that separates the lowest 25% from the upper 75%.

The norm-referenced approach can be applied at the state or school level. If applied at the state level, the same potential drawback described above – variable group size by school – will be an issue. If applied at the school level, the size of the equity group will be proportionately consistent for all schools (e.g. 25% of the students). However, the characteristics of that group could vary quite broadly. For example, in a very high performing school the lowest 25% may in fact represent students above proficient and not necessarily at-risk. At a very low performing school many at-risk students may not be included in the equity subgroup.

Ultimately, it is a policy decision regarding whether to use a norm-referenced threshold and, if so, at what level it should be applied (i.e. state, district, or school). However, given the discussions of the AAC and the strong priority the committee placed on inclusion of all at-risk students in the equity approach, *it is likely more reasonable to consider either a criterion-referenced approach or a norm-referenced approach defined at the state-level*.

Selecting a Metric

Elementary and Middle Schools

Continued use of the current NSPF growth to achievement gap metric in the elementary and middle schools remains a strong option. Specifically, the school level MGP would be evaluated against the AGP and the resulting percent meeting AGP is the school's equity metric. This has worked well in NSPF previously, is familiar and well-understood by many stakeholders, and offers the advantage of linking the equity outcome to specific attainment target.

High Schools

For high school, the current approach subtracts the percent proficient and above performance of the focal group (i.e., either ELL, SPED or FRL-eligible) from the percent proficient and above performance achieved by the reference group or all students. Although this current approach may be appealing in that the information communicated appears to be simple, there are some technical challenges associated with the design. For example, if the vast majority of students in the reference group (all students) are also in the focal group (either an IEP, FRL, OR ELL student), the lack of a difference could be attributed to the fact that both reference and focal groups consist largely of the same set of students. Another issue for consideration is that if a school consistently has a small super subgroup (e.g., meets the minimum n size requirement of 10), having just one or two students in the subgroup fall short of proficiency could drastically change the relationship and the size of the perceived gap between the much larger reference and focal group.

Possible considerations for NDE to explore to find optimal solutions for challenges with the current approach used in high school:

- In lieu of using "all students" as a reference group, consider setting a standard aligned with the AMO or other meaningful target. Use this standard as a benchmark against which the subgroup performance is evaluated.
- Consider increasing the minimum n-size to ensure that the performance of just one or two students would not drastically increase the size of the gap between reference and focal groups. One strategy may entail having to consolidate students across years to ensure that smaller schools can still be evaluated on this indicator.

College and Career-Readiness

The AAC reviewed each of the four CCR indicators currently in the framework and provided several recommendations and considerations for re-evaluating this component. Presently, the four indicators that comprise CCR on the NSPF are: 1) Percentage of students receiving remediation in higher education; 2) Percentage of students earning an advanced diploma from high school; 3) Percentage of students earning an Advanced Placement or International Baccalaureate exam credit or at least one college credit from a higher education institution through post-secondary opportunities or an articulated CTE program; and, 4) Percentage of students participating in the ACT and SAT tests.

Recommendations for modifying current indicators

Common recommendations made by the AAC for modifying three out of the four current indicators in the CCR component are:

- NDE should consider revising the percentage of advanced diploma earners indicator to include all high school diplomas earned by students.

Several AAC members (with no members dissenting) recommended that this indicator needs to be modified for CCR. These AAC members questioned whether an advanced diploma should be given higher value than other diplomas. The AAC members who advocated for amending this indicator noted that unless there is evidence that an advanced diploma yields better post-secondary outcomes for students than other high school diplomas, identifying advanced diplomas as the only diploma to be considered under this indicator is questionable.

- Substitute the percentage of students participating on ACT or SAT tests with the percentage of students meeting college-ready benchmarks on either test.

Since all students will be required to take either the ACT or SAT tests, the AAC agreed that this indicator no longer provides a useful signal of how many students in a given high school are academically ready for post-secondary studies. According to AAC members, the college-readiness benchmarks used for each test would provide a good indicator for evaluating how

many students at each high school can meet basic admission requirements at post-secondary institutions. Several AAC members (with no dissenting members) also noted that including the readiness signal from college-readiness benchmarks would render the remediation rates indicator unnecessary to include in the framework since the college-readiness benchmarks would yield a more comprehensive picture of which students are likely to require remediation. The college-readiness benchmarks provide a more comprehensive picture since remediation rates are only computed based on the sub-set of students who move into the NV higher education system but are not tracked for those attending institutions outside of the system. In contrast, the college-readiness benchmark information will apply to all students in the system.

Additional Recommendations to CCR

The AAC noted that all of the current indicators only focus on college readiness and that a career readiness focus is missing in this component. Several suggestions provided by the AAC members included:

- Explore the feasibility of using some of the non-academic skills outlined in David Conley's framework. According to the small group that made this suggestion (no opposing comments from other members), this exploration should help determine to what extent these types of measures should be included for accountability. The recommendation was made that these indicators may be reported but not used for accountability.

- Offer other CCR indicator options to account for inequities in resources across schools to offer AP or IB courses.

- Move the credit-earning indicator currently located in the "Other Indicator" category to CCR (either as another indicator or as an option to AP or IB). Credit-earning should also be extended to grade 10.

- One group noted the importance of reporting CCR indicators by subgroup performance (no opposing comments from other members).

AAC members noted that CCR should be weighted more in the framework only if career-readiness measures are available. However, if career-readiness is not included, some AAC

members noted a preference for lowering the current weight attributed to the CCR component or leaving the weight at its current level.

CCR Design Illustrations

Although indicators for CCR vary across states, Glancy et al. (2014) reported that some of the common indicators found across 28 states using CCR indicators include many of the indicators included in the current NSPF. Per Glancy et al., common indicators found across accountability frameworks developed by 28 states include:

- Dual enrollment participation and/or completion
- Advanced Placement/International Baccalaureate participation and/or results
- Graduation rates
- ACT/SAT participation and/or results
- Postsecondary participation rate
- Industry certifications earned
- College remediation rates

As reflected by the set of common indicators above, the majority of states use indicators that largely reflect college readiness and do not include non-academic focused skills that are becoming increasingly viewed by educational researchers as CCR indicators. Considering that there is scant information and research on the use of these indicators for accountability in most states, this underscores the point made by one group of AAC members that the inclusion of indicators addressing non-academic focused skills will likely need to be explored and tested prior to using for accountability.

Ideas for non-academic focused skills and career readiness can be located in several resources (e.g., Conley, 2012; Del Razo, et al., 2014). For example, Del Razo et al. at the Annenberg Institute for School Reform note that "In addition to acquiring content knowledge, students need to be able to work and share ideas with others, establish and maintain positive relationships, lead, think creatively…and achieve positive goals for themselves…students need these skills and abilities to be able to apply their knowledge and continue their learning beyond high school graduation (p. 31). Based on this premise, Del Razo et al. recommend developing indicators where students demonstrate mastery in the following areas:

- Communication skills
- Leadership skills
- Technology skills
- Relationship building
- Civic involvement
- Creative expression

However, defining metrics for these indicators would require setting competency standards to determine "mastery." Additionally, no reference is made by the authors to point to the set of instruments that can be used to collect these data. Should these types of indicators be

considered, the feasibility of data collection and technical considerations on the appropriate uses of these indicators will need to be evaluated by a CCR committee and/or the accountability TAC.

In reference to suggestions made by the AAC to modify current indicators, Table 5 presents an illustration of one possible way to re-configure this component:

Table 5. Possible Reconfiguration CCR based on AAC Input

Current NSPF	Revised Version	Reported (not for accountability)
Remediation Participation rates in ACT or SAT	Percentage of students meeting college-readiness benchmarks for ACT or SAT	Disaggregated data by subgroups
AP/IB credit or college credit	Option for schools to select either one: Grade 9 and 10 on-track for credit-earning; AP/IB credit earned from exam; industry certification.	Disaggregated data by subgroups
Percentage of students earning advanced diploma	Percentage of students earning traditional diploma OR consider reporting graduation rates.	Disaggregated data by subgroups

Missing from this illustration is the inclusion of additional career readiness indictors that go beyond industry certification. Although Glancy et al. (2014) suggest using assessments such as WorkKeys or considering passing scores on industry certification or licensure exams to assess career-readiness, any new assessment or scores considered for the NSPF will require further evaluation.

School Climate

Guidance on School Climate

Although AAC members did not initially agree to the inclusion of this component to the framework, consensus was reached after specific priorities and guidance were outlined for considering School Climate. The specific guidance provided by AAC members for shaping this new component for the framework was as follows:

- School climate serves as a "celebration" of what schools do for their students. All schools can earn full points on school climate but may fall short (i.e., lose points on this component) if sub-group gap issues emerge from possible school climate indicators (see next bullet).
- Alternatively, a conjunctive "take all" or "lose all" approach can be used for school climate indicators. If a school meets standards set across all indicators on school climate, they receive full points for this component on the framework, if not, then they forfeit all school climate points.
- School climate indicators should help spotlight whether there is an over-representation of minorities who are being disciplined (e.g., suspensions, expulsions, etc.).
- All indicators must have clear metrics and can be measured with quality instruments.

- There should be flexibility provided to ensure that schools can select indicators that work best for their schools. However, this flexibility should still meet the condition stated in the bullet above.

The most critical guidance point provided by the AAC is that school climate should not be punitive. Hence, all AAC members agreed that school climate should have a moderate to minimal impact on the framework. However, there was unanimous agreement that indicators that are common to all schools for School Climate could potentially function in a more punitive manner especially if these may uncover disproportionate disciplinary rates impacting traditionally disadvantaged students (e.g., African Americans). Based on the guidance provided, three categories for identifying school climate indicators emerged. The first category would consist of indicators common to all school such as discipline data, and attendance data. The second category would consist of indicators where schools have flexibility to select different indicators. This second category would provide an opportunity for schools to select indicators that would show-case the good work they do to improve school quality for teachers, students, parents, and the larger community. The third area serves as an optional "test-site" where schools can choose to collect and report results from these indicators but not use the information for accountability.

Two areas where no clear consensus was achieved were the use of school climate survey results (e.g., teacher, parent and student perception about school climate surveys) and the inclusion of teacher turnover data as an indicator in one of the 3 categories established. In reference to the use of school climate survey results, three AAC members thought this should be used for category 3 since results reported can be highly influenced by principals collecting the data. All of the district-based AAC members thought survey results could be included under category 2 since districts can co-develop surveys together to ensure the development of a quality survey. One member thought this could be used as a category 1 indicator since there is a common survey instrument used in districts where schools can be compared on questions related to climate.

School Climate Illustration Based on AAC Guidance

Based on the guidance received from AAC members, Table 6 presents an illustration of possible indicators falling into each of the three categories for school climate. The inclusion of this component would replace the "Other Indicator" category in the current NSPF with average daily attendance becoming one indicator located in category 1. For high school, as noted in the CCR section, the 9th grade credit earning that is also currently located under "Other Indicator" would be moved to CCR and also augmented to include grade 10. We recommend the reporting of school climate survey results under category 3 until the results from these surveys are deemed suitable by NDE and other stakeholder groups to use as either a category 1 or 2 indicator.

Table 6. Illustration of possible indicators for School Climate

Category 1 – Common or standardized state-wide school climate indicators	Category 2 – District-specific school climate indicators for schools	Category 3 – Reported school climate information
• Student discipline	Criteria and guidance should	• Teacher, student and

(e.g., expulsions, suspensions) • Safe learning environment (e.g., bullying, dangerous incidents reported) • Attendance	be developed by a school climate committee for districts to submit evidence for Category 2. A menu of indicators that may be similar across districts could be developed for this category. Criteria to include: • Must be measurable using clear metrics. • Should use instruments that meet quality standards. • Should reflect data that has provided useful information to the school and district.	parent perceptions about school climate • Teacher turnover data • Leadership turnover data • Percentage of parents attending parent-teacher conferences. • Percentage of teachers working on school activities outside of regular school hours

An important area of consideration for NDE is that data collection efforts for school climate indicators would require additional resources to invest in suitable instruments and undertake quality assurance checks. Furthermore, the data reported under category 3 should be accompanied by caveats to forewarn the audience that these data need to be understood in context. For example, if a school has a lower percentage of parents attending parent-teacher conferences relative to other schools, this finding may not be an indication of "weaker" school climate, but rather could be reflective of other contextual factors that should get highlighted (e.g., a school serving over-age and under-credit youth).

Design Decisions

The AAC was also asked to provide recommendations about overall model design decisions. We address those recommendations in this section focusing on: combining multiple measures, weighting, flexibility, and conjunctive criteria.

Combining Multiple Measures

There are at least four approaches to combine multiple indicators to yield a single outcome: *compensatory*, *conjunctive*, *disjunctive*, and *profile* methods. Compensatory means that higher performance in one measure may offset or compensate for lower performance on another measure. Conjunctive means that acceptable performance must be achieved for every measure. Disjunctive means that performance must be acceptable on at least one measure. A profile refers to a defined pattern of performance that is judged to be satisfactory, unsatisfactory, or equivalent. A profile approach is often operationalized using a matrix to combine indicators for making judgments.

The current NSPF is based on a compensatory approach and the AAC affirmed that this method should be retained, except as addressed subsequently in *conjunctive criteria*. The compensatory approach is a technically sound decision as it yields a higher degree of reliability compared to some alternative methods, because the overall decision is based on multiple indicators evaluated more holistically. Moreover, reliability improves because random error in multiple measures tends to cancel.

The compensatory approach is currently operationalized as an 'index' or weighted combination of points which sum to a maximum value of 100. The AAC recommended staying with this approach and we also recommend continuing with this approach as it is straightforward to understanding and operationalize. The point values for each component reflect the weights or values attributed to each category and are discussed in the following section.

Weighting

Decisions about weighting components in the framework should be based on policy values. For this reason, the AAC was asked to review the current weights in the NSPF and provide feedback on any adjustments or changes that should be applied. The AAC did not recommend specific values; however, they did provide general guidance to inform the department's decision making. Those recommendations are noted below:

Elementary and Middle Schools

- Growth should have more influence than status (also reflected in the design principles). The AAC was equally split over whether the weight of growth should be increased beyond the current NSPF emphasis or stay the same. The AAC did agree that if the weight of growth is increased, a modest increase is advisable.
- The weight for achievement gaps or equity should be increased. The AAC acknowledged the importance of equity as a prominent component of the model and the need to closely monitor achievement gaps.
- Attendance was NOT seen as a reliable or important component. Therefore, the AAC was unanimous in recommending that attendance be included as one of several indicators under school climate. However, the weight of school climate may need to be lower than was previously assigned to attendance. As noted previously in this document, school climate should have a nominal influence in the model.

High Schools

- The AAC recommends lowering the total weight assigned to growth s. This decision was influenced by the AAC's recognition that little is known about the availability of growth data in high school.
- A majority of the AAC recommended increasing the weight for achievement gaps to emphasize performance on gaps. Others on the AAC reasoned that gaps are also accounted for under graduation rates, which removes the need to increase the weight.
- A majority recommend ensuring that readiness is weighted more than graduation rate, if career readiness is incorporated into CCR.

- The AAC was unanimous in recommending an increase in weight for 9th grade credit sufficiency/deficiency and that this indicator should be moved under CCR. Several AAC members also recommended incorporation of 10th grade credit sufficiency/deficiency.
- Attendance was NOT seen as a reliable or important component. Therefore, the AAC was unanimous in recommending that attendance be included as one of several indicators under school climate. However, the weight of school climate may need to be lower than was previously assigned to attendance. As noted previously in this document, school climate should have a nominal influence in the model.

Flexibility

As the design principles make clear, the AAC was unanimous in their recommendation that school accountability outcomes should be comparable from school to school within year and across years. This was an important guiding decision and influenced many of the recommendations outlined in this document, which reflect the emphasis on standardized, comparable information.

However, the AAC was unanimous in identifying one area where some local flexibility should be permitted: school climate. As noted in the preceding section, the recommendation for the school climate component consists of three categories, two of which combine to exert nominal influence on accountability outcomes. One of those two categories influencing ratings would include district specific school climate indicators. By so doing, the AAC reasoned, some latitude will be permitted for inclusion of measures thought to be more authentic and relevant to the diverse systems and schools, without compromising comparability of outcomes overall, due to the negligible influence of this category.

The AAC also discussed the potential for exploring flexibility for district assessments for purposes such as measuring achievement gaps, measuring academic growth, or more broadly reflecting CCR. However, the AAC did not have the time to pursue these ideas more fully and acknowledged it is premature to consider these approaches for near-term, high-stakes implementation.

Conjunctive Criteria

Finally, the AAC offered a recommendation for two areas that should not be fully compensatory: high school graduation and achievement gaps. Specifically, the committee acknowledged that there should be a minimum threshold of performance for these indicators, below which schools may not receive the most favorable ratings. Some illustrative options for operationalizing this recommendation are:

- Any school that falls below the target graduation rate or achievement gap score will have the overall rating decreased by one category.
- Multiple regions will be identified below the target graduation rate and achievement gap score, each of which corresponds to progressively severe reductions in the overall rating. For example, performance in the lowest of three categories results in a three rating reduction; performance in the second lowest results in a two rating reduction and so forth.

In addition to one of the above recommendations, it is recommended that specific supports and inventions apply to any school failing to meet the minimum threshold in graduation or achievement gap.

Since the AAC did not set specific values for the weights and thresholds assigned to each indicator and component, we recommend that this be addressed at the standard setting meeting. In the next section we outline analyses for evaluating the efficacy of recommendations made by the AAC. These analyses should be conducted to determine the viability of recommendations put forward by the AAC or to determine whether alternative methods to address the recommendations need to be explored or further refined prior to operationalizing.

Analyses to Evaluate Efficacy

For each of the indicators, components, and the overall ratings assigned to schools, there are some technical areas to check with desired features and principles articulated by the AAC. In this section, we outline suggested analyses to check to see whether desirable results are achieved for the corresponding features.

Monitoring Equity across Components and Indicators

Given the AAC's emphasis on ensuring the equity category includes at-risk groups, including racial/ethnic minorities, the state should produce descriptive analyses to report the percent of students in these groups included in the performance-based consolidated group. It is clear that the department regularly tracks this information and, in fact, provided numerous analyses during the course of AAC meetings to show group membership prominently reflected in target subgroups when a performance based criterion is applied. However, these equity reports should also be expanded to monitor how different groups perform on the CCR indicators and the School Climate indicators. Low performance found in schools with historically underperforming groups will contribute to decisions on "triggers" or business rules to ensure that these schools cannot attain the highest ratings on different components or overall ratings.

Monitoring the Impact of Contextual Factors

To mitigate concern that the selected approach systematically advantages or disadvantages schools based on the demographics of the students served or by factors such as school size, the department should produce analyses showing the range of scores by key demographic factors known to impact results, such as a school's enrollment size, percent FRL or percent of racial/ethnic minorities served. This can be easily displayed as a scatterplot with points earned on the y-axis and the demographic indicator (e.g. percent FRL) on the x-axis. A strong approach that conforms to the desired design principles will not reveal a strong positive correlation and will yield a full range of results, including favorable results, for all schools, including those that serve a large percentage of at-risk students.

Monitoring Stability of Results

An approach that is more stable from year to year is desired over an approach that is highly variable. For this reason, comparing the school level equity score correlations across years for multiple approaches is advisable. This can be accomplished by retroactively computing the equity score that would have been earned had the selected methodology been applied in 2013 and 2014 and comparing the stability with that of the NSPF gap score (or another competing approach) in those same years.

Reviewing Contributions of Indicators and Components on Overall Ratings

The weights attributed to the components in the system play a critical role in determining the effective contribution of each piece to the overall rating achieved by a given school. The AAC recommended a general weighting scheme to ensure that components of lesser importance (e.g., the component formerly called "other indicators") contribute substantially less to a school's overall rating than other components such as performance gaps, growth and achievement to ensure schools stay focused on improving academic performance. Since specific weights were not recommended by the AAC, the NDE will need to test out results from assigning different weights to each component to ensure that a focus on academic performance is achieved with enough emphasis placed on the other features of the system.

Conclusion and Next Steps

This report serves to document the recommendations from the AAC to guide the development of Nevada's new school accountability system. Moving forward, we suggest the following two phase process to incorporate the recommendations in the revised operational system. The recommended process outlined below takes into account the challenge presented by the unexpected loss of summative assessment data in 2015.

Phase I: Analyze Data and Develop Model Revisions (Summer 2015 – Fall 2015)

Based on the recommendations summarized in this report, we suggest the NDE move forward with various analyses to evaluate the viability of implementing the recommendations made by the AAC. In some cases, the analyses will rely on data from 2014 due to limited assessment data in 2015. The analyses conducted will include comparison of results using specific business rules for suggested "triggers" relative to gaps, considering different weighting schema for the different components of the NSPF and evaluating outcomes relative to using different approaches for evaluating sub-group performance. We further understand that Nevada's Technical Advisory Group (TAG) will convene to review analyses conducted and provide their input on design choices. This adds to the credibility and defensibility of the process and we strongly support the TAG's engagement.

The culmination of the process should be the development a revised Nevada School Performance Framework. Due to incomplete assessment data in 2015, incorporating growth in a manner consistent with the framework will not be possible for the 2016 implementation. For this

reason, the 2016 model will be 'transitional' and the first implementation of the complete model that includes growth will occur in 2017.

Following completion of the analyses and development of specific decision rules for the revised model, we recommend reconvening the AAC to review the proposal. Currently, the AAC is set to meet in September, 2015; although given the scope of the recommendations, it may be appropriate to reschedule that meeting to a later date.

Phase II: Standard Setting (Summer 2016)

The first set of complete statewide Smarter Balanced Assessment data will not be available until following the 2016 administration. For this reason, we recommend delaying standard setting to the summer of 2016 when the impact of academic achievement measures can be evaluated and incorporated in the standard setting process. This will enable the process to utilize school profiles that accurately portray performance, which is critical to the process. At this time, we propose to set standards for the transitional system and estimate standards for the complete system. Because the estimated standards will be based on incomplete growth data, it will be prudent to conduct an additional review in the summer of 2017 before confirming the standards. However, we believe it is reasonable to take advantage of the 2016 standard setting meeting to establish 'candidate' standards, pending confirmation or revision in 2017.

References

Conley, D. (2012). Measuring college and career readiness by means of proficiency or competency-based systems. In *The State of College Access and Completion: Improving College Success for Students from Underrepresented Groups*. New York: Routledge.Darling-Hammond, L. (2006). No child left behind and high school reform. *Harvard Educational Review*, *76*(4), 642-667.

Del Razo, J. L., Saunders, M., Renée, M., López, R. M., & Ullucci, K. (2014). Leveraging Time for School Equity: Indicators to Measure More and Better Learning Time. *Annenberg Institute for School Reform at Brown University*.

Glancy, E., Fulton, M., Anderson, L., Zinth, J. D., Millard, M. & Millard, M. Blueprint for College Readiness (Denver, CO: Education Commission of the States, October 2014) Retrieved from: http://www.ecs.org/docs/BlueprintforCollegeReadiness

Novak, J. R., & Fuller, B. (2004). Penalizing Diverse Schools? Similar test scores, but different students, bring federal sanctions. In *Saving Our Schools: The Case for Public Education: Saying No to" No Child Left Behind"*, Berkeley, CA: RDR Books.

Attachment A

Accountability Advisory Committee Members

Member	Position
Arnold, Leslie	Assistant Superintendent, Clark County School District (CCSD) Supt.
Burckhard, Bridget	Treasurer, Nevada Parent Teacher Association
Combs, Allison	Assistant Vice Chancellor Nevada System of Higher Education
Cranor, Erin	President, CCSD Board of Trustees District G
Demchak, MaryAnn	Professor of Special Education & Graduate Director University of Nevada Reno. Representing the Nevada Special Education Advisory Committee
DuFrene, Denise	Principal Corbett Elementary School, Washoe County School District
Gavin, Patrick	Director State Public Charter School Authority
Holmes-Sutton, Tonia	CCSD Project Coordinator
Jensen, Dave	Superintendent Humboldt County School District
Lee, Suzette	Chair of Board for Communities in Schools Nevada
Moradkhan, Paul –	Vice President of Government Affairs, Las Vegas Metro Chamber of Commerce
Murilo, Ruben	President of Nevada State Education Association (NSEA)
Salas, Rachel	Assistant Professor of Literacy at University of Nevada, Reno. Representing the Nevada English Mastery Council
Williams, Yvette	Founder of Clark County Black Caucus